Journal de la Société Française de Statistique
Vol. 153 No. 1 (2012)

Infinite-Dimensional Autoregressive Systems and the Generalized Dynamic Factor Model

Titre : Systèmes Autorégressifs de Dimension Infinie et le Modèle à Facteurs Dynamiques Généralisé

Marco Lippi[1]

Abstract: The Generalized Dynamic Factor Model is usually represented as an infinite-dimensional moving average of the common shocks plus idiosyncratic components. Here I argue that such representation can only be the solution (reduced form) of a deeper, structural, infinite-dimensional system of equations which contain both an autoregressive and a moving average part. I give conditions for the solutions of such infinite-dimensional systems to make sense and be weakly stationary, and study their decomposition into common and idiosyncratic components. Interesting links to long memory as generated by aggregation are also shown.

Résumé : Le Modèle à Facteurs Dynamiques Généralisé est habituellement représenté comme une moyenne mobile infinie des chocs communs à laquelle s'ajoutent des composantes idiosyncratiques. Dans cet article, nous expliquons que cette représentation n'est que la solution (sous forme réduite) d'un système d'équations plus profond, structural, de dimension infinie, et contenant à la fois une partie autorégressive et une partie à moyenne mobile. Nous livrons des conditions qui garantissent que les solutions de tels systèmes sont faiblement stationnaires et nous analysons leur décomposition en composantes communes et idiosyncratiques. Enfin, nous établissons des liens avec la mémoire longue qui est engendrée par agrégation.

Keywords: Infinite-Dimensional Datasets, Generalized Dynamic Factor Models, Ensembles de données de dimension infinie, Modèle à Facteurs Dynamiques Généralisé
AMS 2000 subject classifications: 62P20, Applications to Economics

1. Introduction

Consider an infinite-dimensional stochastic process

$$\mathbf{x}_t = (x_{1t}\ x_{2t}\ \cdots\ x_{nt}\ \cdots)'.$$

Suppose that $\mathbf{x}_t$ is zero-mean and weakly stationary, precisely that $\mathbf{x}_{nt} = (x_{1t}\ x_{2t}\ \cdots\ x_{nt})'$ is zero-mean and weakly stationary for all n. Let $\mathbf{\Sigma}_n^x(\theta)$ be the spectral density matrix of $\mathbf{x}_{nt}$, $\theta \in [-\pi\ \pi]$, and let $\lambda_{nj}^x(\theta)$ be the jth eigenvalue of $\mathbf{\Sigma}_n^x(\theta)$. Assume that there exists an integer $q \geq 0$ such that (I) $\lambda_{n,q+1}^x(\theta) \leq D$, for some $D > 0$, for all n and θ almost everywhere in $[-\pi\ \pi]$, and (II) if $q > 0$ then $\lambda_{nq}^x(\theta) \to \infty$ as $n \to \infty$ for θ almost everywhere in $[-\pi\ \pi]$. Then, as Forni and Lippi (2001) prove, there exists a q-dimensional white noise $\mathbf{u}_t = (u_{1t}\ u_{2t}\ \cdots\ u_{q_t})'$, square-summable filters $b_{ij}(L)$, $i \in \mathbb{N}$, $j = 1, 2, \ldots, q$, and an infinite-dimensional, zero-mean, weakly stationary process

$$\boldsymbol{\xi}_t = (\xi_{1t}\ \xi_{2t}\ \cdots\ \xi_{nt}\ \cdots)',$$

[1] Universitá di Roma La Sapienza and Einaudi Institute for Economics and Finance (EIEF)
E-mail: ml@lippi.ws

such that

$$\begin{aligned} x_{it} &= \chi_{it} + \xi_{it} \\ \chi_{it} &= a_{i1}(L)u_{1t} + a_{i2}(L) + \cdots + a_{iq}(L)u_{qt} + \xi_{it} = \mathbf{a}_i(L)\mathbf{u}_t + \xi_{it}, \end{aligned} \tag{1}$$

where:
(i) There exists $B > 0$ such that for all n and θ almost everywhere in $[-\pi\ \pi]$ $\lambda^{\xi}_{n1} \leq B$.
(ii) $\lambda^{\chi}_{nq}(\theta) \to \infty$ as $n \to \infty$, for θ almost everywhere in $[-\pi\ \pi]$. Note that $\mathbf{\Sigma}^{\chi}_n(\theta)$ has rank q for all θ, so that $\lambda^{\chi}_{n,q+s}(\theta) = 0$ for all $s > 0$.
(iii) u_{jt} and $\xi_{i,t-k}$ are orthogonal for all $k \in \mathbb{Z}$, $i \in \mathbb{N}$ and $j = 1, 2, \ldots, q$.

The converse is also true, i.e. if $\mathbf{x}_t$ has a representation of the form (1) with $\boldsymbol{\chi}_t$ and $\boldsymbol{\xi}_t$ fulfilling (i), (ii) and (iii), then the above conditions on the qth and $(q+1)$th eigenvalues of $\mathbf{\Sigma}^{x}_n(\theta)$ hold. Forni and Lippi (2001) also show that the decomposition of x_{it} into χ_{it} and ξ_{it} is unique. Of course neither $\mathbf{u}_t$ nor the filters $b_{ij}(L)$ are unique. For, a transformation of $\mathbf{u}_t$, by an invertible matrix, and the corresponding transformation of the filters produce a representation equivalent to (1). We assume that the white noise $\mathbf{u}_t$ has been normalized, i.e. that $\mathbf{u}_t$ is orthonormal.

Model (1), under (i), (ii) and (iii), is known as a Generalized Dynamic Factor Model. It has been the object of a quite vast literature starting with Forni, Hallin, Lippi and Reichlin (2000), Forni and Lippi (2001), Stock and Watson (2002a and b), Bai and Ng (2002), Bai (2003).

We refer to (i) by saying that the variables ξ_{it} are *weakly correlated* or that they are *idiosyncratic*. The variables u_{jt} are called the *common shocks* and the variables χ_{it} the *common components*. Thus in (1) an infinite dimensional, large-dimensional in empirical applications, vector is driven by a small-dimensional vector of common shocks plus the weakly correlated idiosyncratic terms.

If the variables x_{it} are macroeconomic variables, as in the large majority of empirical applications so far, the shocks u_{jt}, $j = 1, 2, \ldots, q$, represent common macroeconomic sources of variation. The variables ξ_{it} represent instead variable-specific, sectoral, local disturbances or measurement errors. If further information, economic-theory statements in particular, is available, then the structural common shocks can be identified with he same methods applied in the construction of Structural VARs, i.e. determining a matrix $\mathbf{H}$ such that for the structural shocks we have $\tilde{\mathbf{u}}_t = \mathbf{H}\mathbf{u}_t$, while $\tilde{\mathbf{a}}_i(L) = \mathbf{H}\mathbf{a}_i(L)$ are the structural impulse-response functions (see e.g. Forni, Giannone, Lippi and Reichlin, 2009).

The motivation for the present paper is that representation (1) is in general the moving-average solution of a system of equations in which each variable x_{it} depends on common shocks and idiosyncratic components, but also directly on some of the variables x_{jt}, with or without lags. For example, suppose that the variable x_{1t} is the price of good G. Then x_{1t} depends on the prices of other goods that enter G's production process. Or, suppose that x_{1t} results from a decision rule adopted by an optimizing agent under rational expectations. Then x_{1t} typically depends on (i) lagged values of x_{1t}, (ii) other variables belonging to the system, both current and lagged, (iii) a finite moving average of a vector white noise. See Hansen and Sargent (1980), as a paradigm for this kind of models.

Piecing together (a) the above considerations on the autoregressive links between the variables x_{it} and (b) the distinction between common and idiosyncratic shocks, under mild simplifying assumptions, the structural model would take the following form:

$$\mathbf{C}_0\mathbf{x}_t + \mathbf{C}_1 x_{t-1} + \cdots + \mathbf{C}_p\mathbf{x}_{t-p} = \mathbf{B}_0\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_t + \cdots + \mathbf{B}_q\mathbf{u}_{t-q} + \boldsymbol{\xi}_t, \tag{2}$$

or, in compact version

$$\mathbf{C}(L)\mathbf{x}_t = \mathbf{B}(L)\mathbf{u}_t + \boldsymbol{\xi}_t, \tag{3}$$

where the matrices $\mathbf{C}_j$ have an infinite number of rows and columns, while the matrices $\mathbf{B}_k$ have an infinite number of rows and q columns.

There are cases in which the existence of a stationary solution for system (3) requires nothing more than the usual condition. For example, if the system is

$$x_{it} = \alpha_i x_{1,t-1} + b_i u_t + \xi_{it},$$

then of course a stationary solution exists if and only if $|\alpha_1| < 1$. However, if the system is genuinely infinite dimensional, for example

$$x_{it} = \alpha_i x_{i+1,t-1} + b_i u_t + \xi_{it},$$

then the whole infinite-dimensional matrix $\mathbf{C}(L)$ must be taken into consideration.

The paper studies the following simplification of (3):

$$x_{it} = \sum_{j=1}^{\infty} c_{ij} x_{j,t-1} + b_i u_t + \xi_{it}, \tag{4}$$

or, in compact form,

$$(\mathbf{I} - \mathbf{C}L)\mathbf{x}_t = \mathbf{b}u_t + \boldsymbol{\xi}_t. \tag{5}$$

Of course the finite moving average on the right-hand side of (3) does not play any role for the existence of stationary solutions. Thus having $q = 1$ and $\mathbf{B}(L) = \mathbf{b}$ does not imply any los of generality. On the other hand, generalizing the results proved here to model (3) is a fairly easy task.

Section 2 provides conditions on the infinite-dimensional matrix $\mathbf{C}$ and the infinite-dimensional vector $\mathbf{b}$, such that equation (5) has a stationary solution. I find that if $\mathbf{C}$ is a bounded operator on the Banach space of bounded sequences c_i, the norm being $\sup_i |c_i|$, if $\sup_i |b_i| < \infty$ and the spectral radius of $\mathbf{C}$ is less than unity, then the representation

$$\mathbf{x}_t = (\mathbf{I} - \mathbf{C}L)^{-1}\mathbf{b}u_t + (\mathbf{I} - \mathbf{C}L)^{-1}\boldsymbol{\xi}_t$$

makes sense and is a stationary solution of (5). Section 3 shows that if $\mathbf{C}'$ is a bounded operator on the Hilbert space of square summable sequences d_i, the norm being $\sqrt{\sum_i |d_i|^2}$, and the spectral radius of $\mathbf{C}'$ is less than unity, then $\mathbf{x}_t$ is a Generalized Dynamic Factor Model, with $q = 1$, in which $(\mathbf{I} - \mathbf{C}L)^{-1}\mathbf{b}u_t$ and $(\mathbf{I} - \mathbf{C}L)^{-1}\boldsymbol{\xi}_t$ are the common and the idiosyncratic component respectively. In Section 4 I discuss some examples and links to long-memory stochastic processes. Section 5 concludes.

2. Existence of a stationary solution

It will be convenient to re-index the variables x_{it} in $\mathbb{Z}$ instead of $\mathbb{N}$, so that (4) becomes

$$x_{it} = \sum_{j=-\infty}^{\infty} c_{ij} x_{j,t-1} + b_i u_t + \xi_{it}. \tag{4}$$

Assumption 1. *The sequence ξ_{it}, $i \in \mathbb{Z}$, is zero-mean, weakly stationary and idiosyncratic. Moreover, u_t is a unit-variance white noise.*

Assumption 2. $\lim_{n\to\infty} \sum_{j=-n}^{n} b_j^2 = \infty$.

Assumption 3. $u_t \perp \xi_{j,t-k}$ *for $i \in \mathbb{Z}$ and $k \in \mathbb{Z}$.*

Assumption 2 implies that the first eigenvalue of the spectral density matrix of

$$(b_{-n} \; \cdots \; b_0 \cdots \; b_n)' u_t$$

tends to infinity as $n \to \infty$. As a consequence, under Assumptions 1, 2 and 3, defining $y_{it} = b_i u_t + \xi_{it}$, the sequence y_{it} is a Generalized Dynamic Factor Model with $q = 1$ (note that the dynamics for the variables y_t may only come form the idiosyncratic terms).

As an example, we may think of an infinite production system, whith c_{ij} and b_j representing, respectively, the quantity of commodity j and labour necessary to produce one unit of commodity i. If x_{it} and u_t are the the price of commodity i and the wage rate respectvely, then (4) is the price equation, with ξ_{it} being a disturbance term, under the assumption of a zero rate of profit (or of a factor profit embodied in the coefficients c_{ij}). Given the non-negative vector

$$\mathbf{d} = (\cdots \; d_{-n} \; \cdots \; d_0 \; \cdots \; d_n \; \cdots),$$

the equation

$$(\mathbf{I} - \mathbf{C}')\mathbf{z} = \mathbf{d}$$

determines the activity levels necessary to obtain **d** as a net product. Special cases are

$$x_{it} = c_{ii} x_{i,t-1} + b_i u_t + \xi_{it}, \tag{6}$$

in which each industry employs only its own product as a means of production, and

$$x_{it} = c_{i,i-1} x_{i-1,t-1} + b_i u_t + \xi_{it}, \tag{7}$$

in which the only mean of production used in industry i its the product of industry $i-1$. Further discussion of these cases is postponed to Section 5.

An obvious candidate for a solution of equation (5) is

$$\mathbf{x}_t = (\mathbf{I} + \mathbf{C}L + \mathbf{C}^2 L^2 + \cdots)\mathbf{b} u_t + (\mathbf{I} + \mathbf{C}L + \mathbf{C}^2 L^2 + \cdots)\boldsymbol{\xi}_t. \tag{8}$$

But of course we must impose conditions on the infinite-dimensional matrix **C** and the vector **b**. It seems natural to assume that $b_i < b$ for some positive b. Moreover, we must have that all the components of **Cb** are finite, i.e. that $\sum_{j=-\infty}^{\infty} c_{ij} b_j$ is finite for all i.

Denote by ℓ_∞ the Banach space of all complex bounded sequences

$$\mathbf{c} = (\cdots \; c_{-n} \; \cdots \; c_0 \; \cdots \; c_n \cdots)',$$

with norm $\|\mathbf{c}\|_{\ell_\infty} = \sup_i |c_i|$.

Assumption 4. *The vector* **b** *belongs to ℓ_∞ and the matrix* **C** *is a bounded operator on ℓ_∞.*

Let us recall that $\mathbf{C}$ is a bounded operator on ℓ_∞ if

$$\sup_{\|\mathbf{d}\|_{\ell_\infty}=1} \|\mathbf{Cd}\|_{\ell_\infty} < \infty,$$

and that the norm of a bounded operator on ℓ_∞ is defined as the left-hand side of the above inequality:

$$\|\mathbf{C}\|_{\ell_\infty} = \sup_{\|\mathbf{d}\|_{\ell_\infty}=1} \|\mathbf{Cd}\|_{\ell_\infty}. \tag{9}$$

Moreover, it easily seen that

$$\|\mathbf{C}\|_{\ell_\infty} = \sup_i \sum_{j=-\infty}^{\infty} |c_{ij}|.$$

Under Assumption 4, the *spectrum* of $\mathbf{C}$ is the subset of the complex field $\mathbb{C}$ containing all λs such that $\lambda\mathbf{I} - \mathbf{C}$ has not a bounded inverse. The spectrum of $\mathbf{C}$ is a bounded, closed subset of $\mathbb{C}$. The *spectral radius* of $\mathbf{C}$, denoted by $r_{\ell_\infty}(\mathbf{C})$, is defined as $\sup|\lambda|$, for λ belonging to the spectrum of $\mathbf{C}$.

Assumption 5. $r_{\ell_\infty}(\mathbf{C}) < 1$.

Under Assumption 5 the series

$$\mathbf{I} + \mathbf{C}\mu + \mathbf{C}^2\mu^2 + \cdots$$

converges with respect to the norm (9) for some $\mu > 1$. This implies that $\|\mathbf{C}^n\|_{\ell_\infty} < A\mu^{-k}$ for some $A > 0$ and therefore

$$1 + \|\mathbf{C}\|_{\ell_\infty} + \|\mathbf{C}^2\|_{\ell_\infty} + \cdots < \infty. \tag{10}$$

For the spectrum of bounded operators on Banach spaces and the results mentioned above, see Dunford and Schwartz (1988), p. 566-7, Lemma 4 in particular.

Proposition 1. *Under Assumptions 1 through 5, (8) is a weakly stationary solution of equation (4). Moreover* $\sup_i \mathrm{E}(x_{it}^2) < \infty$.

Proof. I want to prove that the right-hand side of (8) has finite second moments. Consider first the common component of the ith variable in (8):

$$(b_i + \sum_{j=-\infty}^{\infty} c_{ij} b_j L + \sum_{j=-\infty}^{\infty} c_{ij}^{(2)} b_j L^2 + \cdots) u_t,$$

where $c_{ij}^{(k)}$ denotes the (i, j) entry of $\mathbf{C}^k$. The filter's squared gain is

$$\begin{aligned}
|b_i + \sum_{j=-\infty}^{\infty} c_{ij} b_j e^{-i\theta} + \sum_{j=-\infty}^{\infty} c_{ij}^{(2)} b_j e^{-i2\theta} + \cdots|^2 & \\
&\le \|\mathbf{b}\|_{\ell_\infty}^2 (1 + \sum_{j=-\infty}^{\infty} |c_{ij}| + \sum_{j=-\infty}^{\infty} |c_{ij}^{(2)}|^2 + \cdots)^2 \\
&\le \|\mathbf{b}\|_{\ell_\infty}^2 (1 + \|\mathbf{C}\|_{\ell_\infty} + \|\mathbf{C}^2\|_{\ell_\infty} + \cdots)^2.
\end{aligned}$$

Consider now the idiosyncratic component

$$\xi_{it} + \sum_{j=-\infty}^{\infty} c_{ij}\xi_{jt} + \sum_{j=-\infty}^{\infty} c_{ij}^{(2)}\xi_{jt} + \cdots = (\mathbf{I}_{[i]} + \mathbf{C}_{[i]}L + \mathbf{C}_{[i]}^2 L^2 + \cdots)\boldsymbol{\xi}_t$$

where $\mathbf{C}_{[i]}^k$ denotes the ith row of $\mathbf{C}^k$. Then consider

$$(\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}L + \mathbf{C}_{[i,m]}^2 L^2 + \cdots)\boldsymbol{\xi}_t, \tag{11}$$

where $\mathbf{C}_{[i,m]}^k$ denotes the row vector

$$(c_{i,i-m}^{(k)}\ c_{i,i-m+1}^{(k)}\ \cdots\ c_{i,i}^{(k)}\ \cdots\ c_{i,i+m-1}^{(k)}\ c_{i,i+m}^{(k)}).$$

For the spectral density of (11) we have

$$\begin{aligned}(\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}e^{-i\theta} + \mathbf{C}_{[i,m]}^2 e^{-i2\theta} + \cdots)\boldsymbol{\Sigma}_{[i,m]}^{\xi}(\theta)(\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}e^{-i\theta} + \mathbf{C}_{[i,m]}^2 e^{-i2\theta} + \cdots)' \\ \leq \|\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}e^{-i\theta} + \mathbf{C}_{[i,m]}^2 e^{-i2\theta} + \cdots\|_{\mathrm{E}}^2 \lambda_{i+m,1}^{\xi}(\theta) \\ \leq \|\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}e^{-i\theta} + \mathbf{C}_{[i,m]}^2 e^{-i2\theta} + \cdots\|_{\mathrm{E}}^2 B,\end{aligned}$$

where $\|\cdot\|_{\mathrm{E}}$ is the standard Euclidean norm, $\boldsymbol{\Sigma}_{[i,m]}^{\xi}(\theta)$ is the spectral density of the vector $(\xi_{i-m,t}\ \cdots \xi_{i+m,t})$, $\lambda_{i+m,1}^{\xi}(\theta)$ is the first eigenvalue of $\boldsymbol{\Sigma}_{i+m}^{\xi}(\theta)$, which is greater or equal to the first eigenvalue of $\boldsymbol{\Sigma}_{[i,m]}^{\xi}(\theta)$ (because $\boldsymbol{\Sigma}_{[i,m]}^{\xi}(\theta)$ is a submatrix of $\boldsymbol{\Sigma}_{i+m}^{\xi}(\theta)$, see Forni and Lippi, [9]). On the other hand,

$$\begin{aligned}\|\mathbf{I}_{[i,m]} + \mathbf{C}_{[i,m]}e^{-i\theta} + \mathbf{C}_{[i,m]}^2 e^{-i2\theta} + \cdots\|^2 &= |c_{i,i-m}e^{-i\theta} + c_{i,i-m}^{(2)}e^{-i2\theta} + \cdots|^2 \\ &+ \cdots + |1 + c_{ii}e^{-i\theta} + c_{ii}^{(2)}e^{-i2\theta} + \cdots|^2 + \cdots + |c_{i,i+m}e^{-i\theta} + c_{i,i+m}^{(2)}e^{-i2\theta} + \cdots|^2 \\ &\leq (|c_{i,i-m}| + |c_{i,i-m}^{(2)}| + \cdots)^2 \\ &+ \cdots + (1 + |c_{ii}| + |c_{ii}^{(2)}| + \cdots)^2 + \cdots + (|c_{i,i+m}| + |c_{i,i+m}^{(2)}| + \cdots)^2 \\ &\leq \left(1 + \sum_{j=-m}^{m} |c_{i,i+j}| + \sum_{j=-m}^{m} |c_{i,i+j}^{(2)}| + \cdots\right)^2 \leq \left(1 + \|\mathbf{C}\|_{\ell_\infty} + \|\mathbf{C}^2\|_{\ell_\infty} + \cdots\right)^2.\end{aligned}$$

The conclusion follows, with $\mathrm{E}(x_{it})^2$ bounded by $(\|\mathbf{b}\|_{\ell_\infty}^2 + B)(1 + \|\mathbf{C}\|_{\ell_\infty} + \|\mathbf{C}^2\|_{\ell_\infty} + \cdots)^2$. Q.E.D.

3. The solution of the autoregressive equation as a Generalized Dynamic Factor Model

Consider a zero-mean weakly-stationary infinite dimensional process η_{it}, $i \in \mathbb{Z}$. According to condition (i), see the Introduction, $\boldsymbol{\eta}_t$ is idiosyncratic if $\lambda_{n1}^{\eta}(\theta)$ is essentially bounded. Now consider a sequence of infinite-dimensional filters

$$\mathbf{f}_n(L) = (\cdots\ f_{n,-n}(L)\ f_{n,-n+1}(L) \cdots\ f_{n,0}(L) \cdots\ f_{n,n-1}(L)\ f_{n,n}(L)\ \cdots), \tag{12}$$

and assume that

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \sum_{j=--\infty}^{\infty} |f_{n,j}(e^{-i\theta})|^2 d\theta. \tag{13}$$

The most obvious example is the arithmetic average:

$$f_{n,j}(L) = \begin{cases} 0 \text{ if } |j| > |n| \\ (1+2n)^{-1} \text{ if } |j| \leq |n|. \end{cases}$$

Forni and Lippi (2001) prove that the sequence $\boldsymbol{\eta}_t$ is idiosyncratic if and only if (13) implies that

$$\lim_{n\to\infty} \mathbf{f}_n(L)\boldsymbol{\eta}_t = 0, \tag{14}$$

in mean square.

We denote by ℓ_2 the Hilbert space of all complex square-summable sequences

$$\mathbf{d} = (\cdots\ d_{-n}\ \cdots\ d_0\ \cdots\ d_n\ \cdots)',$$

with norm $\|\mathbf{d}\|_{\ell_2} = \sqrt{\sum_{i=-\infty}^{\infty} |d_i|^2}$.

Assumption 6. *The matrix $\mathbf{C}'$ is a bounded operator on ℓ_2, i.e.*

$$\sup_{\|\mathbf{d}\|_{\ell_2}=1} \|\mathbf{C}'\mathbf{d}\|_{\ell_2}^2 < \infty.$$

The ℓ_2-norm is defined as $\|\mathbf{C}'\|_{\ell_2} = \sqrt{\sup_{\|\mathbf{d}\|_{\ell_2}=1} \mathbf{C}'\mathbf{d}}$, the spectrum and the spectral radius are defined as above, $r_{\ell_2}(\cdot)$ denoting the spectral radius. We suppose that

$$r_{\ell_2}(\mathbf{C}') < 1.$$

Let $\mathscr{L}_2$ be the Hilbert space of all sequences

$$\mathbf{f} = (\cdots\ f_{-n}\ \cdots\ f_0\ \cdots\ f_n\ \cdots)',$$

where

(1) f_n is defined on $[-\pi\ \pi]$, takes values in $\mathbb{C}$ and is square integrable;

(2) $\int_{-\pi}^{\pi} \sum_{j=-\infty}^{\infty} |f_j(\theta)|^2 d\theta < \infty$.

The $\mathscr{L}_2$-norm is defined as

$$\|\mathbf{f}\|_{\mathscr{L}_2} = \sqrt{\int_{-\pi}^{\pi} \sum_{j=-\infty}^{\infty} |f_j(\theta)|^2 d\theta} = \sqrt{\int_{-\pi}^{\pi} \|\mathbf{f}(\theta)\|_{\ell_2}^2 d\theta}.$$

Now consider the function $\mathscr{C}$, with sends $\theta \in [-\pi\ \pi]$ on the matrix $\mathbf{C}'e^{-i\theta}$. For $\mathbf{f} \in \mathscr{L}_2$, using Assumption 6:

$$\|\mathscr{C}^k\mathbf{f}\|^2 = \int_{-\pi}^{\pi} \|\mathbf{C}'^k e^{-ik\theta}\mathbf{f}(\theta)\|_{\ell_2}^2 d\theta \leq \|\mathbf{C}'^k e^{-ik\theta}\|_{\ell_2}^2 \|\mathbf{f}\|_{\mathscr{L}_2}^2 = \|\mathbf{C}'^k\|_{\ell_2}^2 \|\mathbf{f}\|_{\mathscr{L}_2}^2,$$

so that $\mathscr{C}^k$ is a bounded operator on $\mathscr{L}_2$ and $\|\mathscr{C}^k\|_{\mathscr{L}_2} \leq \|\mathbf{C}'^k\|_{\ell_2}$. Moreover, for $n < m$,

$$\begin{aligned} \|\mathscr{C}^n + \mathscr{C}^{n+1} + \cdots + \mathscr{C}^m\|_{\mathscr{L}_2} &\leq \|\mathscr{C}^n\|_{\mathscr{L}_2} + \|\mathscr{C}^{n+1}\|_{\mathscr{L}_2} + \cdots + \|\mathscr{C}^m\|_{\mathscr{L}_2} \\ &\leq \|\mathbf{C}'^n\|_{\ell_2} + \|\mathbf{C}'^{n+1}\|_{\ell_2} + \cdots + \|\mathbf{C}'^m\|_{\ell_2}. \end{aligned}$$

Under Assumption 6, (10) holds in the ℓ_2-norm. Therefore, denoting by $\mathscr{I}$ the identity operator on $\mathscr{L}_2$,

$$\mathscr{I}+\mathscr{C}+\mathscr{C}^2+\cdots$$

converges in the $\mathscr{L}_2$-norm. Thus, in conclusion, $\mathscr{I}-\mathscr{C}$ is invertible and

$$(\mathscr{I}-\mathscr{C})^{-1}=\mathscr{I}+\mathscr{C}+\mathscr{C}^2+\cdots$$

Proposition 2. *Under Assumptions 1 through 6 the infinite-dimensional process* $\mathbf{x}_t$*, as defined in (8), is a Generalized Dynamic Factor Model with* $q=1$*, where* $\boldsymbol{\chi}_t^*=(\mathbf{I}+\mathbf{C}L+\mathbf{C}^2L^2+\cdots)\mathbf{b}u_t$ *is the common component and* $\boldsymbol{\xi}_t^*=(\mathbf{I}+\mathbf{C}L+\mathbf{C}L^2+\cdots)\boldsymbol{\xi}_t$ *is the idiosyncratic component.*

PROOF. Given a complex-valued function defined on $[-\pi\ \pi]$, we denote by $\underline{f}(L)$ the filter $\sum_{k=-\infty}^{\infty}a_{f,k}L^k$, where $a_{f,k}$ is the coefficient of $e^{-ik\theta}$ in the Fourier expansion of f. Given $\mathbf{f}\in\mathscr{L}_2$, we define

$$\underline{\mathbf{f}}(L)=(\cdots\underline{f}_{-n}(L)\cdots\ \underline{f}_0(L)\ \cdots\ \underline{f}_n(L)\ \cdots).$$

Note that $\mathbf{f}$ is a column whereas $\underline{\mathbf{f}}(L)$ is a row vector. The component $\boldsymbol{\xi}_t^*$ is idiosyncratic if and only if, given the sequence $\mathbf{f}_n\in\mathscr{L}_2$, $\|\mathbf{f}_n\|^2_{\mathscr{L}_2}\to 0$ implies that $\mathrm{E}(\underline{\mathbf{f}}_n(L)\boldsymbol{\xi}_t^*)^2\to 0$ in mean square. Given $\mathbf{f}_n$, let $\mathbf{g}_n=\left[(\mathscr{I}-\mathscr{C})^{-1}\mathbf{f}_n\right]$. We have

$$\underline{\mathbf{f}}_n(L)\boldsymbol{\xi}_t^*=\underline{\mathbf{g}}_n(L)\boldsymbol{\xi}_t.$$

On the other hand, $(\mathscr{I}-\mathscr{C})^{-1}$ is linear, bounded and therefore continuous, so that $\|\mathbf{f}_n\|^2_{\mathscr{L}_2}\to 0$ implies that $\|\mathbf{g}_n\|^2_{\mathscr{L}_2}\to 0$, which in turn implies that $\mathrm{E}(\underline{\mathbf{g}}_n(L)\boldsymbol{\xi}_t)^2\to 0$.

Regarding the component $\boldsymbol{\chi}_t^*$, take the sequence

$$\mathbf{h}_n(\theta)=\frac{1}{\sum_{j=-n}^{n}b_j^2}(\cdots\ 0\ b_{-n}\ \cdots\ b_0\ \cdots\ b_n\ 0\ \cdots)'$$

and observe that $\|\mathbf{h}_n(\theta)\|_{\ell_2}\to 0$ for all $\theta\in[-\pi\ \pi]$, so that obviously $\mathbf{h}_n\in\mathscr{L}_2$. Moreover,

$$\underline{\mathbf{h}}_n(L)\boldsymbol{\chi}_t=\underline{\mathbf{h}}_n(L)\mathbf{b}u_t=u_t.$$

Then define $\mathbf{g}_n=[(\mathscr{I}-\mathscr{C})\mathbf{h}_n]$:

$$\underline{\mathbf{g}}_n(L)\boldsymbol{\chi}_t^*=\underline{\mathbf{h}}_n(L)\boldsymbol{\chi}_t=u_t.$$

Now consider

$$\mathbf{g}_n=(\cdots\ g_{n,-m}\ g_{n,-m+1}\cdots\ g_{n,0}\cdots\ g_{n,m-1}\ g_{n,m}\ \cdots)',$$

and define $\mathbf{g}_{n,s}$, $s>0$, as the truncation of $\mathbf{g}_n$ obtained by setting equal to zero all entries of $\mathbf{g}_n$ whose index exceed s in modulus. There exists a sequence $\mathbf{g}_{n,m_n}$, with $m_n\to\infty$ as $n\to\infty$, such that

$$\underline{\mathbf{g}}_{n,m_n}(L)\boldsymbol{\chi}_t^*\to u_t$$

in mean square. For the spectral density of $\underline{\mathbf{g}}_{n,m_n}(L)\boldsymbol{\chi}_t^*$, we have

$$s_n(\theta)=\mathbf{g}_{n,m_n}(\theta)\boldsymbol{\Sigma}_{m_n}(\theta)\overline{\mathbf{g}_{n,m_n}(\theta)}'\leq\lambda_{m_n,1}(\theta)\|\mathbf{g}_{n,m_n}(\theta)\|^2_{\mathrm{E}}.$$

Now observe that

$$\|\mathbf{g}_n(\theta)\|_{\ell_2} = \|(\mathbf{I}-\mathbf{C}'e^{-i\theta})\mathbf{h}_n(\theta)\|_{\ell_2} \leq \|\mathbf{h}_n(\theta)\|_{\ell_2}(1+\|\mathbf{C}'\|_{\ell_2}) \to 0$$

for all $\theta \in [-\pi\ \pi]$, so that

$$\|g_{n,m_n}(\theta)\|_{\mathrm{E}} \leq \|\mathbf{g}_n(\theta)\|_{\ell_2} \to 0$$

for all $\theta \in [-\pi\ \pi]$. Leting Λ be the subset of $[-\pi\ \pi]$ where $\lambda^{\chi^*}_{n,m_n,1}(\theta)$ does not tend to infinity, $s_n(\theta) \to 0$ in Λ. Because $\underline{\mathbf{g}}_{n,m_n}(L)\boldsymbol{\chi}^*_t \to u_t$ in mean square, we have

$$\int_{-\pi}^{\pi} |s_n(\theta) - (2\pi)^{-1}| d\theta \to 0.$$

This implies that Λ has measure zero. In conclusion, the first eigenvalue of the spectral density of χ_{nt} diverges almost everywhere in $[-\pi\ \pi]$. Q.E.D.

4. Examples. Long memory processes

An important observation is that, unlike finite-dimensional autoregressive equations, in our case the condition that the spectral radius of $\mathbf{C}$, as an operator on ℓ_∞, is sufficient but not necessary for stationarity of $\mathbf{x}_t$. Consider example (6) and suppose that $|c_{ii}| < 1$ but that $\sup_i |c_{ii}| = 1$. In this case 1 belongs to the spectrum of $\mathbf{C}$, i.e. $\mathbf{I}-\mathbf{C}$ has not a bounded inverse. For, $(\mathbf{I}-\mathbf{C})^{-1}\mathbf{v}$, where $\mathbf{v}$ is the vector having unity in all entries, is not a member of ℓ_∞ (note that 1 belongs to the spectrum of $\mathbf{C}$ but has no corresponding eigenvectors). Nevertheless, equations (4) can be solved one after another and each of the variables x_{it} is stationary. Boundedness of $\mathrm{E}(x_{it}^2)$ is not obtained unless we make further assumptions on $\mathbf{b}$ and $\boldsymbol{\xi}_t$. Of course Assumption 6 fails to hold and Proposition 2 holds only under further assumptions on $\boldsymbol{\xi}_t$.

Consider example (7). We have, assuming $b_i = 1$ and using the notation $c_{i,i-1} = \alpha_i$,

$$x_{it} = [u_t + \alpha_i u_{t-1} + \alpha_i\alpha_{i-1}u_{t-2} + \cdots] + [\xi_{it} + \alpha_i\xi_{i-1,t-1} + \alpha_i\alpha_{i-1}\xi_{i-2,t-2} + \cdots].$$

If $\sup_i |\alpha_i| < 1$ for all i, then the spectral radius of both $\mathbf{C}$ and $\mathbf{C}'$ are less than unity and Propositions 1 and 2 hold. But neither $\sup_i |\alpha_i| < 1$ nor $\sup_i |\alpha_i| \leq 1$ is necessary. What matters is the behavior of the terms

$$\gamma_{is} = \prod_{k=0}^{s} \alpha_{i-s},$$

as $s \to \infty$. For instance, if $|\alpha_j| > 1$ only for a finite number of values of j, the remaining ones being bounded away from unity, the spectral radius of $\mathbf{C}$ would be less than unity. On the other hand, under $\sup_i |\alpha_i| = 1$, the process x_{it} can exhibit long memory. Interestingly, here long memory arises for individual variables from the infinite-dimensional autoregressive system, rather than from aggregation as in the literature following Granger (1980), see also Zaffaroni (2004).

Lastly, consider the example

$$x_{it} = \alpha_i x_{0,t-1} + u_t + \xi_{it}, \tag{15}$$

in which all variables depend on $x_{0,t-1}$. Assuming $|\alpha_0 < 1$ we have a stationary solution, while the values of α_i, $i \neq 0$, play no role. However, if for example $\alpha_i = \alpha$ for all $i \neq 0$, then the stationary solution depends on two common shocks, namely u_t and $\xi_{0,t-1}$, so that a common shock arises with the inversion of $\mathbf{I}-\mathbf{C}L$. Clearly here Assumption (6) does not hold.

Systems that are very close to the infinite-dimensional autoregression studied here have been considered in Dupor (1999) and Horvath (2000). These papers study the possibility that autoregressive links among productive sectors cause idiosyncratic components to produce macroeconomic effects. An important difference between this paper and these contributions is that instead of an infinite-dimensional system, their models consist of sequences of non-nested finite-dimensional autoregressive systems which grow in size. My observations above on long memory are in the spirit of their results.

In a fairly different perspective, di Mauro et al. (2007), Pesaran and Chudik (2010) and Chudik and Pesaran (2011) also consider non-nested finite-dimensional sequences of VAR systems. They also assume a sparse structure for the matrix $\mathbf{C}$, according to neighborhood and dominant relationships (example (15) is a case with a dominant sector/country). This makes estimation of the autoregressive coefficients possible. Such a structure is an important difference with respect to the model studied in the present paper and those in Dupor (1999) and Horvath (2000). The aim of the these studies is to determine general conditions under which, for example, long memory or macroeconomic effects emerge from the inversion of $\mathbf{I}-\mathbf{C}L$. However, without further assumptions the autoregressive matrix cannot be estimated.

5. Conclusion

If the spectral radii of $\mathbf{C}$ and $\mathbf{C}'$, as operators on ℓ_∞ and ℓ_2 respectively, are smaller than unity, then system (4) has a stationary solution. Moreover, no common shocks are produced from the idiosyncratic components $\boldsymbol{\xi}_t$ through inversion of $\mathbf{I}-\mathbf{C}L$. However, if the above conditions do not hold then individual variables may exhibit long memory. Moreover, common shocks may arise from inversion of $\mathbf{I}-\mathbf{C}L$.

References

[1] J. Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171, 2003.

[2] J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

[3] A. Chudik and M. H. Pesaran. Infinite-dimensional VARs and factor models. *Journal of Econometrics*, 163(1):4–22, 2011.

[4] F. di Mauro, L. V. Smith, S. Dees, and M. H. Pesaran. Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22(1):1–38, 2007.

[5] N. Dunford and J. T. Schwartz. *Linear Operators, Part I.* Wiley, New York, 1988.

[6] B. Dupor. Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43(2):391–409, 1999.

[7] M. Forni, D. Giannone, M. Lippi, and L. Reichlin. Opening the black box: Structural factor models with large cross sections. *Econometric Theory*, 25(05):1319–1347, 2009.

[8] M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82(4):540–554, 2000.

[9] M. Forni and M. Lippi. The generalized dynamic factor model: Representation theory. *Econometric Theory*, 17(06):1113–1141, 2001.

[10] C. W. J. Granger. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238, 1980.

[11] L. P. Hansen and T. J. Sargent. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, 2(1):7–46, May 1980.

[12] M. Horvath. Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45(1):69–106, 2000.

[13] M.H. Pesaran and A. Chudik. Econometric analysis of high dimensional vars featuring a dominant unit. Cambridge Working Papers in Economics 1024, Faculty of Economics, University of Cambridge, May 2010.

[14] J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–62, 2002a.

[15] J. H. Stock and M. W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002b.

[16] P. Zaffaroni. Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics*, 120(1):75–102, 2004.